



## King's Research Portal

DOI:

[10.1080/01615440.2019.1684859](https://doi.org/10.1080/01615440.2019.1684859)

*Document Version*

Peer reviewed version

[Link to publication record in King's Research Portal](#)

*Citation for published version (APA):*

Blanke, T., & Aradau, C. (2019). Computational genealogy: Continuities and discontinuities in the political rhetoric of US presidents. *HISTORICAL METHODS*. <https://doi.org/10.1080/01615440.2019.1684859>

### **Citing this paper**

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### **General rights**

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### **Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

## **Computational Genealogy – Continuities and Discontinuities in the Political Rhetoric of US Presidents**

Articulations of discontinuity and moments of dissent have been central to critical historical work. However, such vocabularies and analyses of historical change have received less attention in the emerging field of digital methods. Digital methods based on discerning patterns have focused on continuities, while discontinuities and ruptures have been derivative of trends and patterns. By contrast, genealogical methods attend to the entanglement of continuity and discontinuity, and focus on contingency and singularity. This article proposes to develop methods of computational genealogy to analyse multiple temporalities in historical discourses. We experiment with our proposed computational genealogy using the archive of Inaugural speeches by US Presidents. In particular, we show that there is neither a linear advance to Trump's rhetoric nor an exceptional rupture. Our analysis shows that Trump's speech is much more the struggle of the Republicans with their own past ideas than struggles with Democrats.

Keywords: Computational Genealogy; Inaugurals; Digital Methods; Anomaly Detection; Influence Analysis

## Introduction

In January 2017, the 45th president of the United States, Donald J. Trump, gave his inauguration speech. After a controversial campaign that left a country divided about the new president, the speech did not reassure commentators that Trump would become more presidential in office and reunite the country. This feeling was captured in the alleged comment by former president George W. Bush, who was overheard saying “[t]hat was some weird s...t” (Ali 2017). While inauguration speeches are normally seen as attempts to overcome differences, which had been exaggerated in an election campaign, and to emphasize unity, Trump did not follow this script. One of the most memorable phrases in Trump’s inauguration speech was “American carnage” (Fahrenthold et al. 2017), describing a country in crisis with “abandoned factories” and “rising crime”.

The *Los Angeles Times* called the speech “angry”, “dark”, “aggrieved”, and “reminiscent of the apocalyptic portrait” (Barabak 2017). This tone was even more surprising as Trump did not inherit an immediate major economic or political crisis. Trump’s speech has often been compared to Obama’s first inauguration address in 2009, which took place directly after the near complete collapse of the global financial and economic system in 2008. Nevertheless, Obama 2009 was much more positive and forward-looking in the spirit of the campaign slogan of “Yes, we can”. The speech was so positive that the *New Republic* wondered whether it should not have been more “bleak” (Fairbanks 2009), given the historical moment.

This paper will analyse Obama’s and Trump’s speeches in the context of all other inauguration speeches before them. This kind of rhetorical analysis is

currently dominated by a new field called culturomics (Michel et al., 2011). Culturomics stands for the investigation of long historical trends through the quantitative analysis of large-scale cultural records and continuities they express. What is interesting about the positioning of the Obama and Trump speeches is the hybrid of continuity and discontinuity. How is discontinuity to be understood and analysed methodologically? For Trump's and Obama's contemporaries, discontinuity only emerges against taken-for-granted assumptions about the unity of 'tradition' of Inaugural speeches or the continuity of 'economic patterns' of growth/crisis.

We propose 'computational genealogy' as an approach that can go beyond the focus on continuities and provide new computational methods to analyse discontinuities. Inspired by Michel Foucault's uses of genealogy for a 'history of the present', we propose computational genealogy as a critical methodology for digital historical research. While critical research in the humanities has emphasized the need to attend to singularities and specificities, digital methods have tended to work with long-term trends and patterns rather than breaks, discontinuities or contingency. This article develops a digital methodology which takes discontinuities in historical analysis seriously.

Historical comparisons using computational techniques have largely focused on patterns and trends. This is the methodological background of the recent debate sparked by the publication of the *History Manifesto* (Guldi and Armitage, 2014). As big data and "digitally enabled research" have brought back 'big' history questions, Guldi and Armitage argue for a "longue-durée synthesis of policy trends on a worldwide scale" (Guldi and Armitage 2014, 91). While they agree that digitally enabled research can "pursue particular moments of

dissent, schism, and utopianism”, their focus is on “a computer-guided timeline of the relative prominence of ideas” (Guldi and Armitage 2014, 91). Guldi and Armitage see big data and digital methods as re-enabling *long durée* research and resisting short-termist methodology in historical research. However, the opposition of short-term/long-term alone does not account for the methodological challenges of continuity/discontinuity or linearity/contingency in historical research.

We present two novel ways of analysing discontinuities for historical research. Firstly, we draw on a computational method to detect anomalies in a rhetorical timeline. Secondly, we introduce the influence plot for a detailed investigation. Such plots are commonly used to describe outlieriness, leverage and overall influence of all the data points in a regression and to clean the analysis from anomalous data points. We shift this analysis around and use the plot to distinguish significant departures from otherwise routine fluctuations in a time series. Thus, we can compare each previous inauguration speech to Obama’s 2009 and Trump’s 2017 addresses. Based on our analysis, we argue that Obama’s speech was more of a break than Trump’s, contrary to what Bush and other contemporaries thought. Trump’s speech appears as a struggle within the Republican party’s narrative and self-understanding.

## **From Culturomics to Computational Genealogy**

Articulations of discontinuity and moments of dissent have been central to critical historical research. However, such vocabularies and analyses of change have received less attention in the emerging field of digital methods (Rogers 2013). Digital methods and modes of digital reasoning within 'cognitive assemblages' (Hayles 2017, 24) are based on discerning patterns and drawing inferences. The spatial and semiotic analysis of patterns has translated into temporal trends and historical continuity rather than rupture and discontinuity. The digital attention to regularities, trends and patterns is reinforced by the use of statistical methods, which have often privileged regularities.

Culturomics is one of the latest instantiations of using large-scale datasets in the "study of human culture" (Michel et al. 2011, 5). A large-scale "quantitative analysis of culture" (Michel et al. 2011) summarises big trends, while changes are discussed as shifts in the meaning of key historical concepts, but not developed in a systematic theory and methodology. For instance, Lansdall-Welfare et al.'s (2017) culturomics approach focuses on trends in historical British newspapers. Highlighting how key concepts of British history change, it does not offer a computational methodology to specifically target these changes. Similarly, in the context of American history, the State of the Union (SoU) addresses, the US presidents' annual addresses have been a popular source for large-scale computational analysis. Using such a corpus that spans centuries, computational approaches have unearthed unique insights such as the importance of the First World War as a reference point for a change in US political rhetoric (Rule et al. 2015).

However, these analyses are focused on long-term trends and tend not to deliver analyses of contingencies, emergence and singular events or what we can call, following Michel Foucault, computational genealogy. Developed by Foucault based on a coinage and initial reflections by Nietzsche (Foucault, 1984), genealogy has become a significant method for critical research. While the relation between genealogy and history has remained a topic of debate, genealogy has been hailed as 'revolutionising history' (Veyne, 1997). Genealogy has been widely understood to attend to singularities, contingency and discontinuity rather than structure, necessity and progress in historical analysis. Foucault formulates genealogy as the relation between discourse and knowledge, understood as a relation of forces and emergences: 'an immense and multiple battle, but not one between knowledge and ignorance, but an immense and multiple battle between knowledges in the plural—knowledges that are in conflict because of their very morphology, because they are in the possession of enemies, and because they have intrinsic power-effects' (Foucault, 2003, 179).

It is beyond the scope of this paper to offer an exhaustive account of genealogy, given the vast literature that has analysed its diverse elements. We propose a genealogical approach to temporality. In Hubert Dreyfus and Paul Rabinow's seminal interpretation of Foucault's oeuvre, genealogy 'seeks out discontinuities where others found continuous development. [...]. [Genealogy] seeks the surfaces of events, small details, minor shifts, and subtle contours.' (Dreyfus and Rabinow, 1983, 106). The association of genealogy with discontinuity and rupture has become the dominant reading of Foucault's methods, to the extent that he was publicly viewed as a philosopher of

discontinuity. However, this interpretation has been challenged as it isolates temporal continuity and discontinuity, repetition and rupture. Dominick LaCapra has argued that Foucault's methodological practice in *History of Madness* entailed 'a concept of temporality in terms of intricate and variable processes of repetition with change' (LaCapra, 2000, 125). Building on LaCapra's insights, Colin Koopman (2013) has most recently proposed multiple temporalities as central to the genealogical endeavour. If the genealogist is attentive to contingency and emergence, this entails methodological attention to historical transformation, which needs to be analysed through multiple temporalities that include both repetition and rupture, continuity and discontinuity. This way, continuity and discontinuity are not isolated from each other but analysed in their entanglements.

Moreover, Foucault does not deny that discontinuity has been present in historical analysis. Rather, his objection is that "the discontinuous was both the given and the unthinkable: the raw material of history, which presented itself in the form of dispersed events - decisions, accidents, initiatives, discoveries; the material, which, through analysis, had to be rearranged, reduced, effaced in order to reveal the continuity of events" (Foucault 2013, 8). Discontinuity appears in variable forms, which need to be analysed empirically rather than assumed to be either derived from trends or revealing a deeper continuity. A large-scale analysis of historical trends in culturomics presents itself as the exact opposite of a genealogical move. Not only does it tend to privilege trends and patterns of continuity, but it isolates discontinuity as a derivation of a trend. We propose that computational genealogy could build on Foucault's



methodological practices of attending to continuities and discontinuities, to singular events as entanglements of repetition with change.

Computational genealogy is not simply about bringing Foucault's genealogy into the world of big data and digital methods. For us, computational genealogy entails three related moves. Firstly, it attends to discontinuity and contingency by tracing multiple temporalities digitally. The *History Manifesto* has been criticised for the equation of "long with significant" (Cohen and Mandler 2015, 535). However, while *long-durée* is equated with 'surveying social change in the aggregate level', micro-history or dissent are not discarded. Only by analysing "vying topics" or invisible archives can the reader "pursue particular moment of dissent, schisms and antagonism" (Cohen and Mandler 2015, 91). Discontinuity become spatialised through topic or archive differentiation. We offer an alternative by tracing outliers and discontinuities computationally and attending to events and emergences within a temporal continuity (or timeline). A computational genealogy would attend to multiple timelines and disruptions and entanglements of discontinuity and continuity rather than large-scale trends and patterns.

Secondly, computational genealogy would share the critical ethos of Foucault's 'history of the present'. In attending to contingent events and transformations, computational genealogy would also challenge taken-for-granted assumptions of our present and who we are. It can thus contribute to "a historical knowledge of struggles and to make use of this knowledge tactically today" (Foucault 1980, 83), thereby destabilising a common perception of computational methods and knowledge as necessarily uncritical. We link

vocabularies from a statistical tradition such as anomalies, outliers, influences, etc. to the critical analysis of historical contingencies, ruptures, etc.

Thirdly, a computational genealogy would attend to the positivity of discourse in Foucault's sense rather than subscribe to a methodological positivism. Foucault's relation to positivism has been a fraught one and he even sometimes embraced the label of a 'positivist'. However, the positivism of genealogy as "gray, meticulous, and patiently documentary" (Foucault, 1984, 76) is not the methodological positivism of observable facts and the scientist as neutral observer. Rather, it needs to be understood in relation to what Foucault has called a positivity 'to designate from afar the tangled mass that I was trying to unravel' (Foucault, 2013, 125). A positivist approach to text would concentrate on regularities expressed in working correlations and larger trends. Instead, we attend to emergence and destabilising elements in the positivity of discourse. While computational methods have been largely used to support the study of patterns and regularities, we propose to render what is marginalised, what is emergent and contingent visible.

To return to Trump's inauguration speech, we propose to analyse it and Obama's first Inaugural as events that are neither linear in a teleological rendition of the rise to power nor completely discontinuous against a supposed background of presidential "tradition". Rather, by turning to the archive of inaugural speeches, we are investigating the temporal multiplicities that constitute a timeline and in particular those events that influence it either way. The inaugural addresses offer a unique snapshot of US history through the words of newly elected presidents. They develop a programme for a presidency, define the political contours of a presidential term and the creation

of an identity in relation to previous speeches. The inaugural speeches create a president's "brand" and specific understanding of the function of government and can thus be indicators of significant political change. They take place only at the beginning of a president's term and often contain promises of change based on campaign promises, as "[p]residents purposefully tie to their brands reflecting specific ideological messages." (Casey 2016, 12).

As the Inaugurals have been widely used for research with a diverse range of computational and quantitative methods, they are furthermore useful to a "little experiment in method", to use Foucault's language (Foucault, 2007, 358). While the Inaugurals promise to provide unique historical insights, this paper is primarily focused on the methodological aspects of computational genealogies.

The first Inaugural was held by George Washington in 1789. The most recent is Donald Trump's from 2016. For our analysis, the inauguration speeches were scraped from a wiki source (Wikipedia 2017), a community data reference site and enriched by including party affiliations, dates, etc. All quotations of inaugural addresses are also taken from the site, unless otherwise indicated.<sup>i</sup> This leads to a relatively small text corpus of 59 speeches, which includes 58 inauguration speeches and one additional exceptional speech by Ford in 1974, to which we return to later. Overall, there are 797,000+ words in the corpus. The shortest speech was given by George Washington in 1793, while the longest by William Henry Harrison in 1841 was almost 50,000 words.

From this corpus, we created three versions of each text. The first one is the raw text version, as downloaded from the web. The second one applies text cleaning such as deleting non-standard characters, removing stop words, extra

whitespaces and punctuation. The third one finally uses stemming and stem-completion to reduce the dimensions of the analysis and to make the results readable for a human. These final two steps reduced the corpus to ~580,000 and ~478,000 tokens respectively. For each of the analyses below, we will indicate which of the three versions of the corpus we use. All texts are length-normalised throughout. Considering the size of the corpus and the differences in the length of the speeches, length-normalisation is, however, no silver bullet. A human reader is required to draw conclusions from the text analysis and eliminate formal elements such as different speech length.

Compared to the examples of culturomics *supra*, our analysis focuses on understanding how the present has come to be and, maybe more importantly, what is missed in the present that could have also been. To describe the different temporalities we also change the viewpoint on time that dominates computational analysis towards a critical understanding of time. Michel et al. (2011) use time as an outside to compare linguistic changes in words that appear *in* it. We fill time by comparing a point in time with the content it contributes to the events of Obama and Trump as well as the content they differ in. As we will see, the Trump and Obama timelines can be easily developed with (Pearson) correlations of word usage, which we map pair-wise over the whole time of the speeches. This reveals a strong linguistic trend in the speeches that is not surprising but poses a great challenge to our analysis, where Trump and Obama are always closer to their contemporaries than those further away from them in the past. We present a new approach that addresses this rhetorical time bias by considering the comparison of inaugural speeches in

history as a time series. Treating the history of Inaugurals as a time series allows us to build on de-trending techniques to identify discontinuities.

## From Periodisation to Timelines

The Inaugurals have been a rich source for computational historical analysis; aptly summarised by Light: “Word-centric techniques turn from the analysis of events to the analysis of text corpora based on patterns in shared content or style.” (Light 2014, 116). Shared patterns reveal, for instance, the use of “power language” and stylistic similarities over time using words that relate to power (Whissell and Sigelman 2001, 255). Common approaches to the Inaugurals have included keyword analysis and clustering of terms. Pre-defined time periods dominate these methods, typical to culturomics. The third approach we demonstrate considers time as an input. Correlations determine the rhetorical timelines of Trump and Obama.

Using pre-determined time intervals, all Inaugurals can be compared to the ones before the end of the American Civil War (1865 inclusive), between the end of American Civil War and the end of the Second World War (1945 inclusive) and after the Second World War until today. Lim (2002) has employed this method of computational content analysis using keywords to describe trends: “My content analysis is directed at the complete set of 264 Inaugurals addresses (...), with the individual word as the unit of analysis ( $N = 1,832,185$ ).” (Lim 2002, 331). Figure 1 is our example of such a keyword analysis, which reveals continuities and discontinuities in the word usages.

In Figure 1, we can see a number of changes in word usage such as the rise of terms such “nation” and “American”, as well as a strong reference to “freedom” and “liberty” after the Second World War. Key political terms such as “peace” disappear after the Second World War. What is most striking is that

“government” is discontinued as a main reference point after the Second World War. The promise of (good) government disappears from the first addresses of newly elected presidents and is especially missing from the 21st century speeches. Other words gain increased currency. Lim notes that “the word democracy appears just twice in the annual messages before 1901 (before the presidency of Theodore Roosevelt, the reputed father of the rhetorical presidency) and 189 times between 1901 and 2000” (Lim 2002, 331). Using a computational rhetorical analysis, Lim describes how “[p]residential rhetoric has become more anti-intellectual, more abstract, more assertive, more democratic, and more conversational” (Lim 2002, 328).

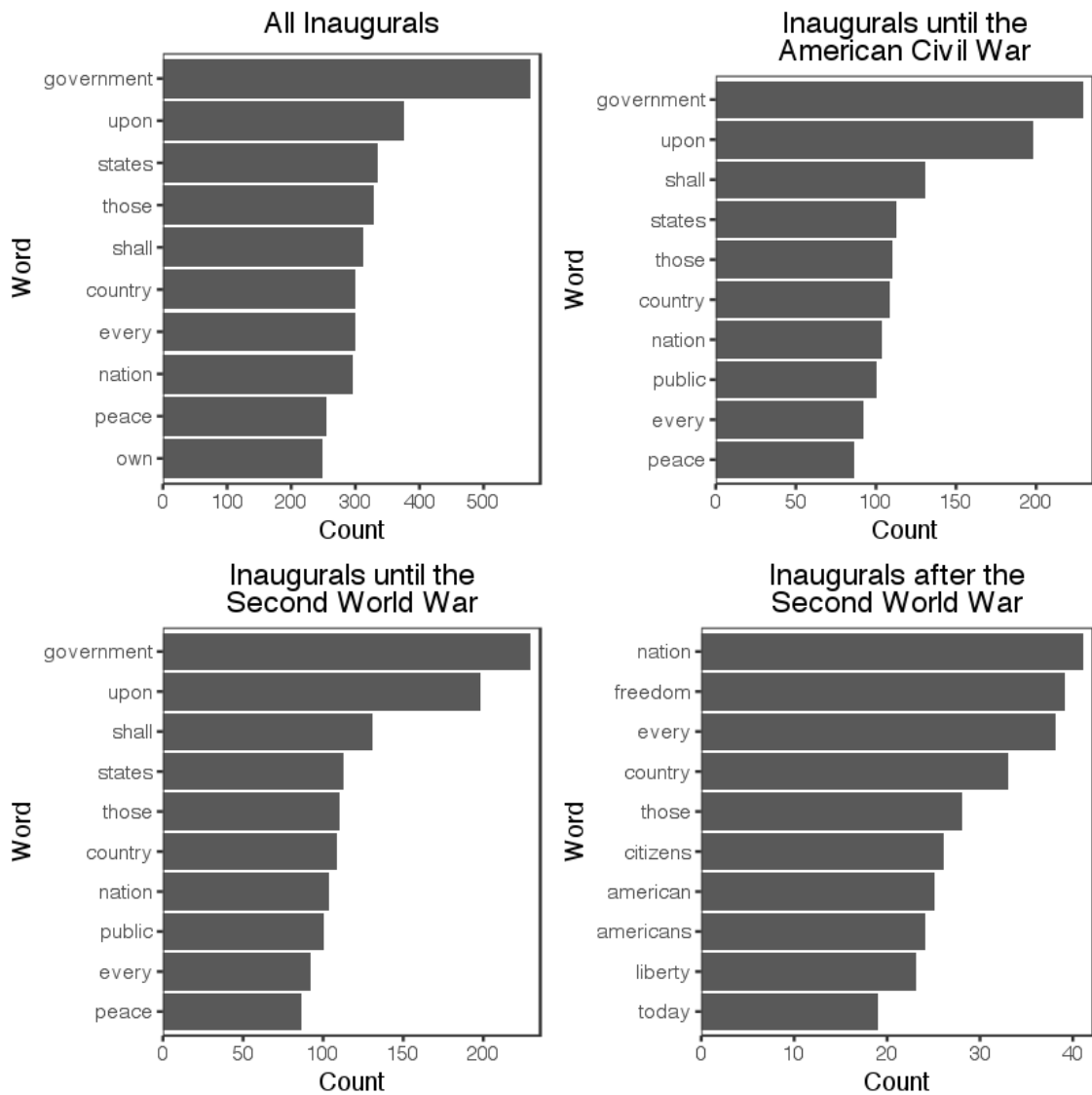


Figure 1: Most Common Words in the Inaugurals

Lim (2002) and our visualisation in Figure 1 are typical historical trend analyses where word continuities and discontinuities are located and compared over specific time intervals that are predefined according to known, major historical events. While this method can lead to original insights, we cannot discover singularities and contingencies in historical narratives, as the time frames are predefined and do not emerge through the analysis. We are interested how such a general trend obscures multiple temporalities that go



beyond known historical fixed points such as wars and what was lost as the trend got consolidated.

Similarly, though more advanced computationally, a typical cluster analysis engages unsupervised machine learning to compare co-occurring words. Cho et al. (2015) have used k-means clusters to discover “trends and patterns” in the Inaugurals until 2010. Among other insights, they discovered a direct relation between the temporal intervals of the Inaugurals and the clusters. Inaugurals in the same cluster are also part of the same time interval, which is something we also observe as a strong temporal bias in the data. Light (2014) uses part-of-speech tagging (PoS) to develop a sophisticated cluster network analysis of the Inaugurals. He creates a “co-word network for the inaugural addresses” using the “147 most prominent (...) words”.

We present hierarchical clustering to visualize continuities and discontinuities. The technique creates a binary tree by successively merging similar speeches, where the more speeches overlap in key terms, the more similar they are. We apply agglomerative clustering on the corpus of stemmed key terms. Speeches and then clusters of speeches are merged until all speeches are merged. Figure 2 visualises the Inaugural clusters. The x-axis depth in the plot is proportional to how dissimilar the children of the subtree are. In the dendrogram, we also colour-coded the N=5 main clusters to provide a better overview.

## Similarity Clusters Dendrogram

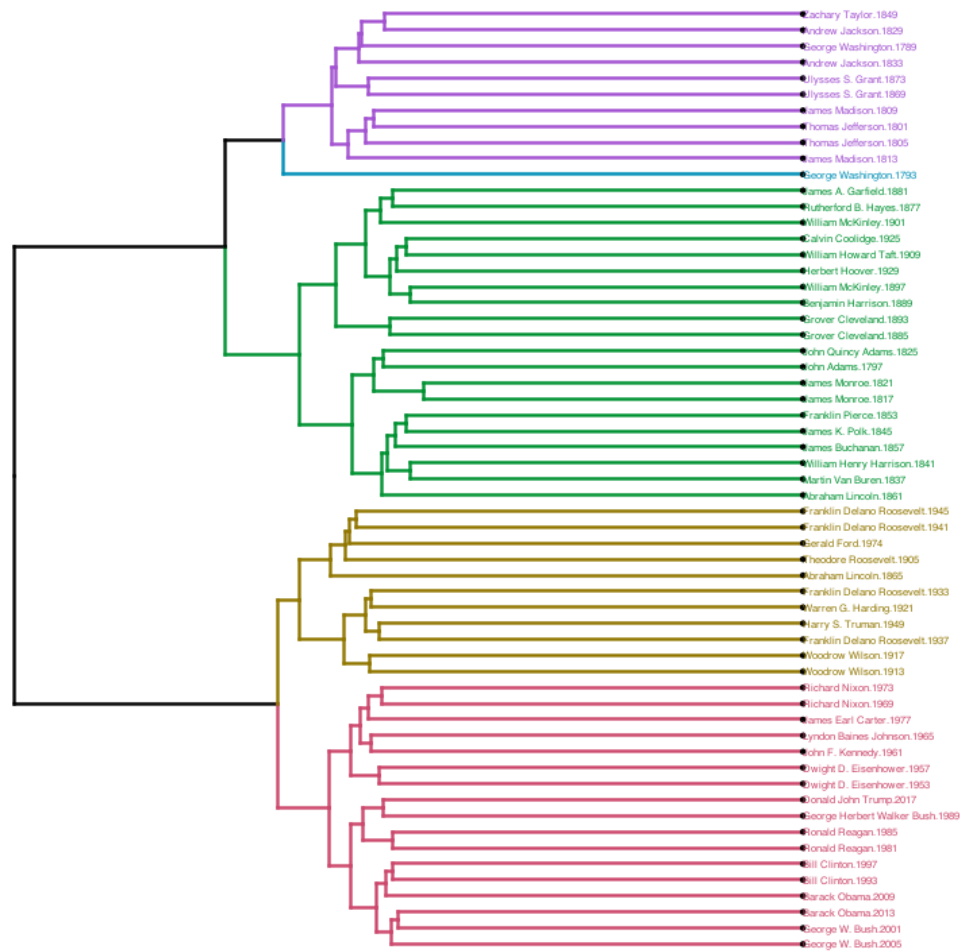


Figure 2: Dendrogram of Inaugural Clusters

The temporal bias in word clusters, as also observed by Cho et al. (2015), is clearly indicated in the results of hierarchical clustering, which clusters those presidents who are close in time (and in person). All five clusters generally follow chronological timelines and known historical events. There is an outlier cluster of George Washington's second speech. We will return to this anomaly further down. Otherwise, there are four clusters, all of which are largely related to known time periods. The first one shows the time period up until the mid 19<sup>th</sup> century. The second temporal cluster lasts until the beginning of the First World

War, the third covers the inter-war period, while the final of the temporal cluster runs from the end of the Second World War and includes the two most recent presidents, Obama and Trump. The most interesting cluster is the inter-war one, a time of rapid change and deep crisis. It contains two unusual additions from outside that period: Gerald Ford from 1974 and Abraham Lincoln second Inaugural from 1865. We will encounter both speeches again during our analysis, as they stand out as points of significant departure in times of crisis. Otherwise, there are two more unusual contributions to the second temporal cluster until the First World War. Coolidge 1925 and Hoover 1929 represent the short period of time between the two world wars where cultural peace and economic growth dominated the political discourse like at the end of the 19<sup>th</sup> century.

The final approach, we encountered in the existing computational analysis of the Inaugurals, targets temporal development and its contingencies beyond pre-defined time periods or broad clusters, as it allows us to focus on timelines by applying a pairwise correlation of the Inaugurals. We used the stemmed and completed corpus with English stop words removed to calculate the Pearson correlation of all the words in Obama 2009 and Trump 2017 compared to all other inaugural speeches. Again, length normalisation is applied. We employ Pearson, as it has also been used in (Light 2014) to analyse the Inaugurals. For all words in two documents  $x$  and  $y$ , the Pearson correlation divides the sum of their frequencies in  $x$  and  $y$  minus their respective mean frequencies  $m(x)$  and  $m(y)$  by the square roots of the squares of  $(x - m(x))$  and  $(y - m(y))$ :  $r = \frac{\sum (x - m(x))(y - m(y))}{\sqrt{\sum (x - m(x))^2 \sum (y - m(y))^2}}$ . Because we only check for words

appearing in two speeches under comparison, there are no negative

correlations. Finally, all results are proportionally scaled to deal with speeches of different lengths.

Table 1 shows the first five correlation entries 706 for Obama 2009 and Trump 2017 in alphabetic order.

	<b>Barack Obama 2009</b>	<b>Donald John Trump 2017</b>
<b>Abraham Lincoln 1861</b>	0.3972430	0.3931778
<b>Abraham Lincoln 1865</b>	0.2937831	0.2016157
<b>Andrew Jackson 1829</b>	0.4055562	0.4044803
<b>Andrew Jackson 1833</b>	0.3459964	0.3431760
<b>Barack Obama 2009</b>	1.0000000	0.6012820

*Table 1: Word Correlations with Obama and Trump*

There is, however, a fundamental problem with this kind of analysis, as the temporal dependency becomes the all-dominating characteristic. Figure 3 demonstrates this dependency, which we need to deal with before we can make sense of different temporalities computationally. Here, we sorted all the correlations according to the times of the speeches and plotted the correlation onto a timeline colour-coding the parties presidents belong to. This shows that Obama (blue) and Trump (red) are both more positively correlated to any speech closer to them in time independent of party affiliations. The trends in Table 1 and Figure 4 are only interrupted by the full correlations of Obama-Obama (2009) and Trump-Trump (2017), which we have replaced for the rest of the analysis with the estimated regression value instead.

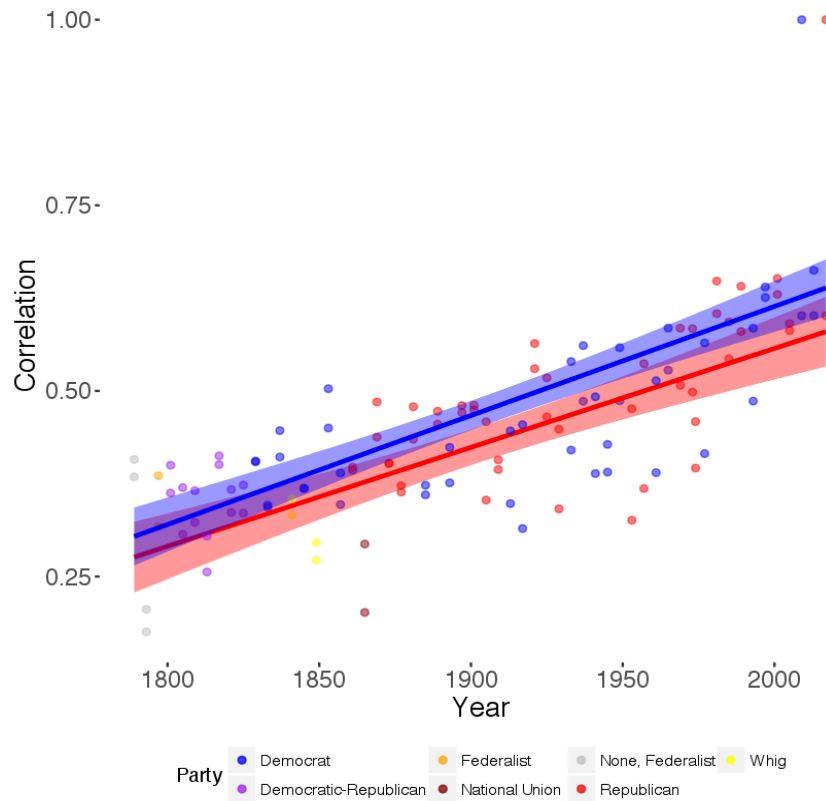


Figure 3: Inaugural Correlations

Mapping correlations into a time series, the temporal ‘bias’ in word usage becomes the all-dominating result. Chronological proximity obscures other temporalities. The closer the speeches are to Obama 2009 (blue) and Trump 2017 (red), the stronger the correlation. Whissell and Sigelman (2001) have also pointed out that time is the overdetermining factor in the computational analysis of the Inaugurals: “Attempts to predict the use of (...) language in inaugural addresses (...) lead to the conclusion that time-based factors are the best predictors of the use of such language (...).”

The regression analysis of similar word usages outputs a temporal relation and has thus the advantage that it does not rely on pre-defined temporal intervals or clusters across time. Thus, we should be able to use it to

find singularities and contingencies. However, this is not easy, as the temporal bias dominates. In order to focus on discontinuities in the regression analysis, we propose to enhance it by addressing time as integral to the speeches. To this end, we employ the methods of time series analysis.

## Problematizing Time – Detrending and Anomalies

To identify events within temporal rhetorical correlations, we need to eliminate the temporal bias in rhetorical timelines. To this end, we conduct a time series analysis of the speech correlations. Figure 3 represents the correlations as a time series; a series of values indexed by years. In time series modeling, we can remove temporal bias with a polynomial model.

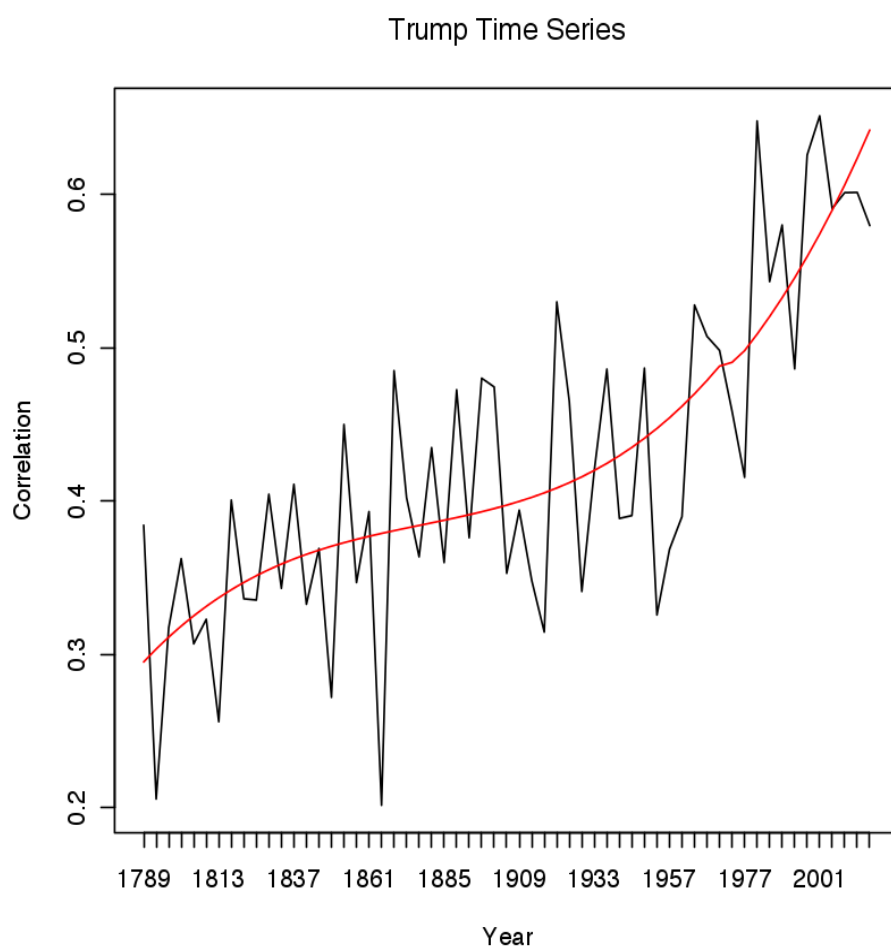


Figure 4: Trump Time Series

Correlations with Trump's 2017 speech constitute a time series, which is visualized in Figure 4. The figure clearly indicates that the time series is not seasonal and exposes the increasing correlation temporal bias. The earlier the

speech, the less related it is to Trump's Inaugural. Figure 5 is the corresponding time series for Obama's 2009 speech, which shows a more linear trend than Trump.

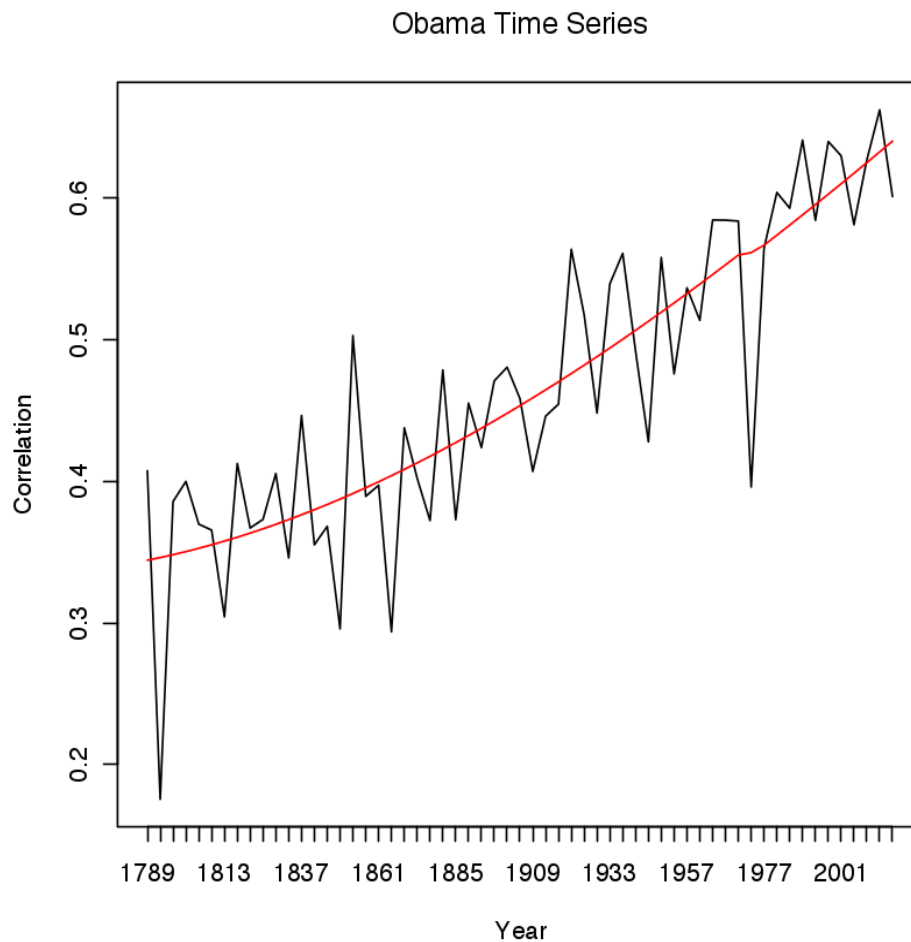


Figure 5: Obama Time Series

Considering correlations as a time series allows us to de-trend the speeches and overcome temporal bias in the Pearson correlations. A trend is a systematic increase or decrease in the series over time, which is not periodic (Cowpertwait and Metcalfe 2009, 5). In the language of time series, the rhetorical timelines of Trump and Obama are both deterministic, as they consistently increase, and they are global, as they apply to the complete series.



In order to de-trend the time series, we construct a polynomial regression model for both Obama and Trump, which is visualized as red lines in Figure 4 and Figure 5. According to the two figures, Obama's model is more linear than Trump's, which is confirmed by the investigation of the polynomial coefficients' significance. For Obama's model, only the first two coefficients (slope and intercept) are significant, while for Trump square(x) is also significant. Obama's model is also a much better fit with R-squared = 0.74, while Trump's model has R-squared = 0.62, which means it is not a great fit but good enough. Only 62% of the variance found in the correlation relation can be explained.

To remove the trend from both time series and de-trend the time series, we subtract the observed values from the regressions' predictions and retrieve the residuals:  $\text{residual}(t) = \text{observation}(t) - \text{prediction}(t)$ . For reasons of space, we only present the de-trended Trump time series in Figure 6. The red regression line is now linear, and the trend removed. Mathematically speaking, the ADF test (Fuller 2009) for the Trump 2017 de-trended time series delivers a p-value of 0.02, while for Obama it is less than 0.01. We consider both time series to be successfully de-trended.

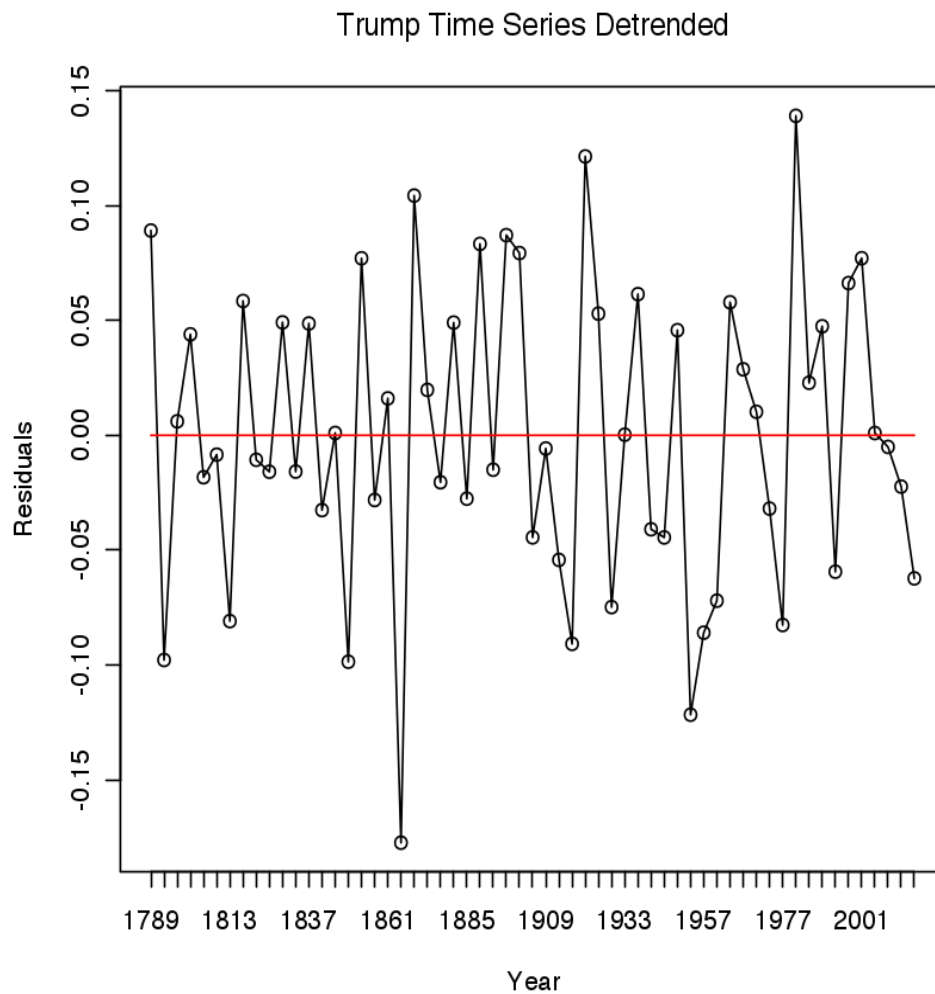


Figure 6: Trump Time Series De-trended

After de-trending, unusual “spikes” appear in the de-trended correlations, which we focus on as moments of discontinuity in the timeline. There is one particular negative spike in the mid 1860s, for instance, which merits further attention. These spikes are points of discontinuity from the Obama and Trump rhetorical timelines. In order to analyse the spikes in the time series, we first apply anomaly detection, which alerts us to these spikes as singular anomalies. With de-trending, we can now concentrate not on the historical continuities the regression exhibits, but on the specificities of several outlying data points.

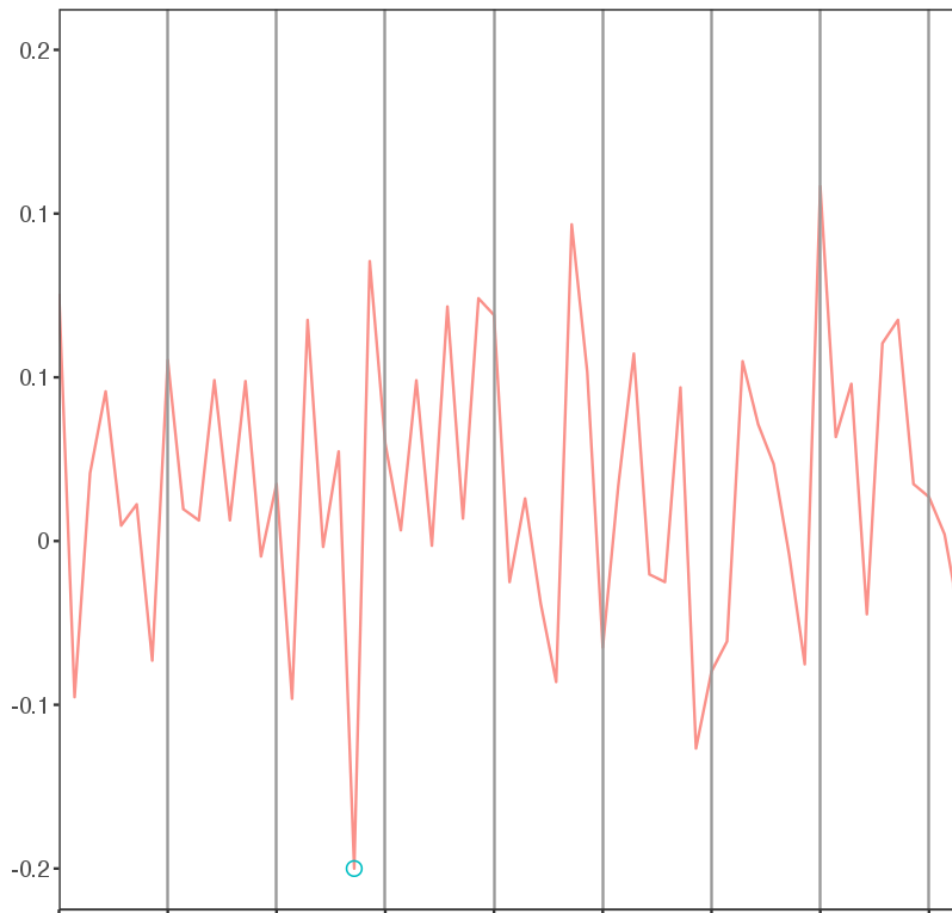
Anomaly detection parses time series data to detect anomalous moments relative to a temporal trend. When analysed in a time series, anomalies are discrepancies or discordances from temporal continuity and linear sequencing (Chandola et al. 2009). These anomalous points can be seen as discontinuous events and singularity. Anomalies in time series are indicated by unexpected spikes or by trend and level changes. We only have one trend, which we removed, and cannot see an abrupt shift in levels. So, we concentrate on distinguishable spikes, which we can identify with a state-of-the-art algorithm presented by (Vallis et al. 2014). It uses the “Seasonal Hybrid ESD Test”, which builds upon the Generalized ESD test (Rosner 1983).<sup>1</sup> The algorithm detects global and local anomalies or those that appear within or outside seasons or periodic fluctuations.

We only show the result of running anomaly detection against Trump’s correlations in Figure 7. With a standard significance level of  $\alpha = 0.05$ , only one anomaly is detected: Lincoln’s 1865 speech. We find the same anomaly for Obama 2009 and also Ford’s 1974 address. While we successfully managed to identify anomalies for Trump and Obama, out-of-the-box anomaly detection is not fine-grained enough. Looking at Figure 7, there are many more interesting spikes, which just miss the required threshold. Next, we turn to a more detailed analysis of these spikes by investigating the so-called ‘influence plot’ of the regressions.

---

<sup>1</sup> The algorithm was originally developed to detect anomalies in the Twitter timeline.

**Trump : 1.69% Anomalies (alpha=0.05, direction=both)**



### Figure 7: Anomalies of Trump's De-trended Time Series

## **Destabilising timelines through Influence Analysis**

To achieve a fine-grained analysis of the “spikes” in the Trump and Obama correlation time series, we turn to influence plots, which can help with a more detailed investigation of our regression models. These plots are often “methods for determining whether a regression model fit to data adequately represents the data” (Fox 2009). Rather than stabilizing the regressions, we are interested in what the influence plot can tell us about the past events that disrupt the general time series and lie beneath the threshold of anomalies we have detected.

An influence plot shows the outlierness, leverage and influence of each data point (Weisberg 2005). In our case, the outliers are those historical speeches that have particular large residuals and are thus furthest away from Obama and Trump. In Figure 8 and Figure 10, the outlierness is represented on the y-axis as residuals, which are normalised relative to leverage (“studentized”). Leverage is measured by the hat-values of the x-axis.

If the hat-value is large than the observation (in our case any speech prior to Obama and Trump) has influenced the prediction a lot. The higher the leverage the further the point will be to the right in the influence plot. The final visual feature in the plots below is the point size, which is proportional to the Cook’s Distance and measures influence. Cook’s Distance is calculated by removing the data point from the model and recalculating the regression (Prabhakaran 2016). Combining leverage and outlierness, an influential event identified by the Cook’s Distance is one which if removed from the data would significantly change the regression.

We plot the regression models used in the de-trending of inauguration speeches and see which data points (speeches) influence Obama's 2009 and Trump's 2017 speeches. Each analysis is furthermore split into two plots that identify the historical intervals and party affiliations respectively. A good spread of both variables demonstrates that influences do not directly depend on either. The numbers in the points are the indexes of the inauguration speeches: no. 1 is the first inauguration by Washington and so on. Obama 2009 is speech no. 57 and Trump 2017 is speech no. 59.

### Obama

The colours in Figure 8 show that neither party affiliation nor period seem to have a strong influence. In both plots, these were included as colouring and are fairly evenly distributed and linked to points of varying size (Cook's distance). In particular, we can see that we can also find outliers closer in time to Obama, while there is still a strong influence from parties that are not the Democrats.

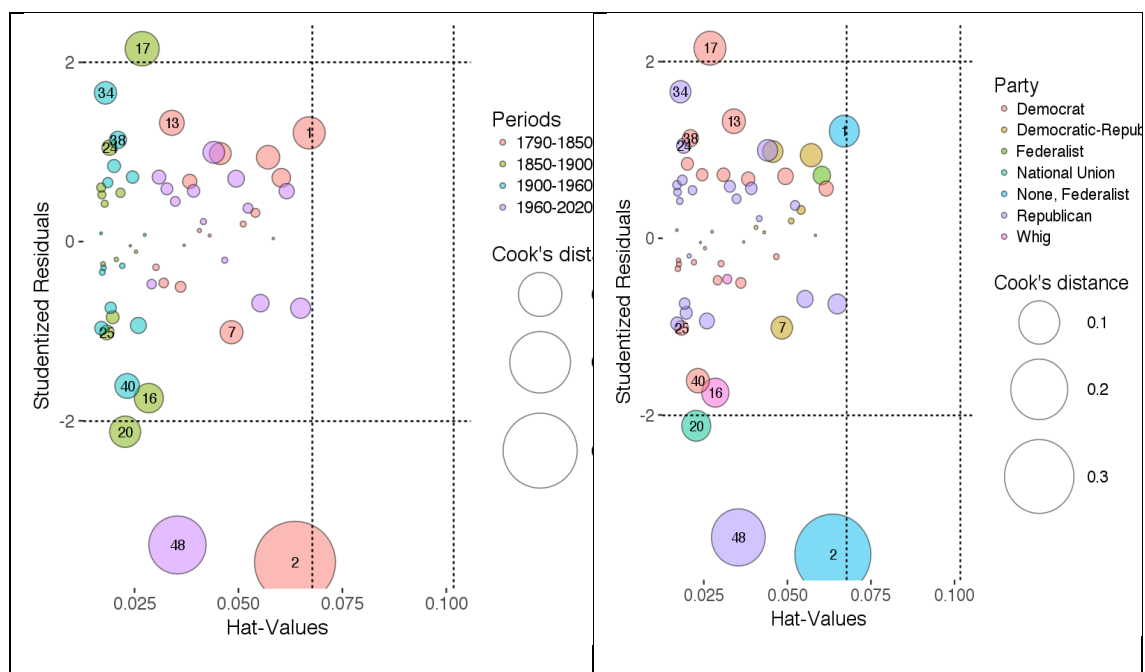


Figure 8: Obama Influence Plot for Parties and Periods

A positive outlier is the 17th inauguration, while the 20th and the 48th speeches are negative ones. The 17th inauguration by Franklin Pierce took place in 1853. The Democrat Pierce actively promoted American overseas influence but in a non-aggressive manner: “The great objects of our pursuit as a people are best to be attained by peace, and are entirely consistent with the tranquillity and interests of the rest of mankind.” (Pierce 1853). While Pierce’s presidency led to divisions and directly preceded the Civil War, his inauguration speech is an attempt to orientate citizens towards global optimism. Overall, however, Obama does not seem to have been strongly positively influenced by any president before him. He really seems to stand for a moment of change.

We have already met the 20th speech as an anomaly. It was Lincoln’s second inaugural address at the end of the American Civil War. It is focussed on the impact of slavery and the civil war. To decode Lincoln’s importance, a good way to further describe its discontinuities with Obama is to focus on respective word differences rather than commonalities or word clusters. Words that distinguish one speech from another allow us to enhance our understanding of what makes speeches stand outside the Obama timeline. To this end, we calculated a table of words that distinguish each Inaugural from all others. Table 2, for instance, shows differences of single terms in Trump’s speech compared to Wilson in 1913 with words unique to Trump, where the words are ordered according to frequency in the speech. The prop column is the proportional importance of the word in the overall corpus of all speeches. We can clearly see that Trump makes distinct references to “American” and “protection”.

word	freq	prop
america	19	0.00505588
american	15	0.01420454
protect	7	0.00374933
back	6	0.00321888

Table 2: Words unique to Donald John Trump 2017 vs Woodrow Wilson 1913

The word variances between Lincoln 1865 and Obama 2009 identify several differences. The most unique words for Lincoln are “offence” and “union” compared to Obama’s “can” and “America”. That Lincoln speaks more frequently about the union and offence should not surprise given the historical circumstances of his speech. Equally, Obama’s most famous reference from his campaign was “yes we can”, and he addresses a unified “America”.

Compared to other speeches, the strongest negative discontinuous speech compared to Obama’s inauguration is the 38<sup>th</sup> speech held in 1974 by Gerald Ford, a Republican who became president after Nixon was impeached. The speech, which Ford denies being an Inaugural, exhibits little confidence in the legitimacy of Ford’s new presidency. It emphasizes continuity to overcome the crisis rather than working for change: “As we bind up the internal wounds of Watergate, more painful and more poisonous than those of foreign wars, let us restore the golden rule to our political process, and let brotherly love purge our hearts of suspicion and of hate” (Ford 1974). There is no confidence in Ford’s leadership, as we find it with Obama. This is further illustrated by the unique terms that distinguish Obama 2009 from Ford 1974. Obama’s distinguishing words are about creating. The three most frequent ones are “new”, “generate”



and “prosper”. The most frequent words unique to Ford 1974 are “constitute”, “prayer” and “elect”, which are indicators of delegated authority. Ford also talks about “burden”, which cannot be found with Obama.

Let us finally take a closer look at the size of the bubbles in the influence plot with a simple plot of Cook’s distances in Figure 9.<sup>ii</sup> The figure does not take into consideration whether the influence is negative or positive but demonstrates the five strongest influences on Obama. We can now see that Washington’s first two speeches influence disproportionately the Obama timeline. Let us not forget that they have set the framework for all future Inaugurals. Their influence should therefore not be a surprise. If we remove these two speeches from the result, we end up with the three most influential speeches, which are in order Ford, Pierce and Lincoln. Ford and Lincoln are the strongest negative moments of departure in the Obama 2009 time series.

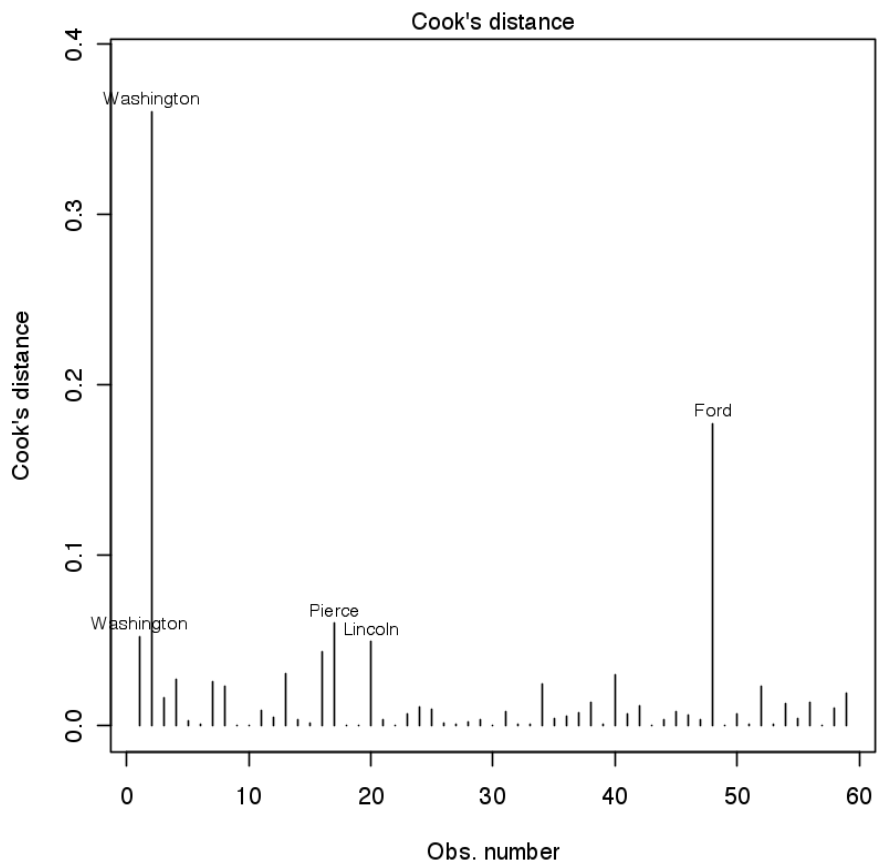


Figure 9: Cook's Distance Obama

### *Trump*

For Trump, almost all points in Figure 10 have either positive or negative residuals. There are very few points in the centre of the residuals, which is not surprising given the fact that the regression is polynomial. Compared to Obama, the Trump time series is more difficult to stabilize. In this sense, he is as unique as Bush's comment seem to indicate. Trump's Cook distances are thus much smaller than for Obama, because the regression is less well-defined. Again, we see no strong relationship to either party affiliation or period. Most notably, for Trump both the strongest positive outliers/discontinuities as well as the most negative ones are Republicans. So, he seems to be involved in a struggle within

the Republican party's own history. Positive outliers are Reagan 1981 (no 50) and Harding 1921 (no 34). Trump has shared comparably more vocabulary with Reagan 1981 than with others, which is not surprising given the early Reagan's influence on all Republican politics after him (Robin 2011). Harding stood for a return to "American normality". During his campaign, Harding declared that America does not need to practice "submergence in internationality, but sustainment in triumphant nationality" (Harding 1920). His speech reflected this strongly nationalist perspective, which is also strongly anti-internationalist. It announces that his election was a referendum by the people on the US retreat from the political world stage after the First World War and the establishment of the League of Nations by his predecessor: "The recorded progress of our Republic, materially and spiritually, in itself proves the wisdom of the inherited policy of non-involvement in Old World affairs. (...). We will accept no responsibility except as our own conscience and judgment, in each instance, may determine." (Harding 1921). Harding's economic plans are also protectionist.

The strongest negative outlier from the Trump timeline is again Lincoln's 1865 speech (no 20). Looking at some of the quotes in the Lincoln, it is fairly easy to see why: "With malice toward none, with charity for all, (...), let us strive on to finish the work we are in, to bind up the nation's wounds, (...), to do all which may achieve and cherish a just and lasting peace among ourselves and with all nations." (Lincoln 1965). In contemporary political analysis, the party of Lincoln is often opposed to the new Republican party of Trump. Thus, this discontinuity should not surprise. Compared to Trump, the earlier Republican

Lincoln speaks frequently about “war” and “offense”. Unique to Trump are strong references to “America”, “countries”, “people” and “protect”.

However, Lincoln is a negative outlier to both Trump and Obama, as the civil war experience is unique. A more interesting discontinuity in the Trump time series is the second most negative outlier, which is another Republican and Eisenhower’s 1953 speech (no 42). After the Second World War, Eisenhower wanted to stand up for the “free of all the world” and speaks of a “common bond”, which “binds the grower of rice in Burma and the planter of wheat in Iowa, the shepherd in southern Italy and the mountaineer in the Andes.” Eisenhower links American freedom with global freedom: “We know, beyond this, that we are linked to all free peoples not merely by a noble idea but by a simple need. No free people can for long cling to any privilege (...)” (Eisenhower 1953). Looking at the unique words for Trump and Eisenhower, Trump compared to Eisenhower uniquely uses “protect” and “back”. Absent from Trump but featuring heavily in Eisenhower’s speech are “faith” and “know”, which indicate an assured path towards an evidence-based future. It is not nationalism that distinguishes Trump’s rhetoric from Eisenhower’s, as both are American patriots, but the belief that this can only be achieved against or with an international order the Republicans themselves set up after the Second World War.

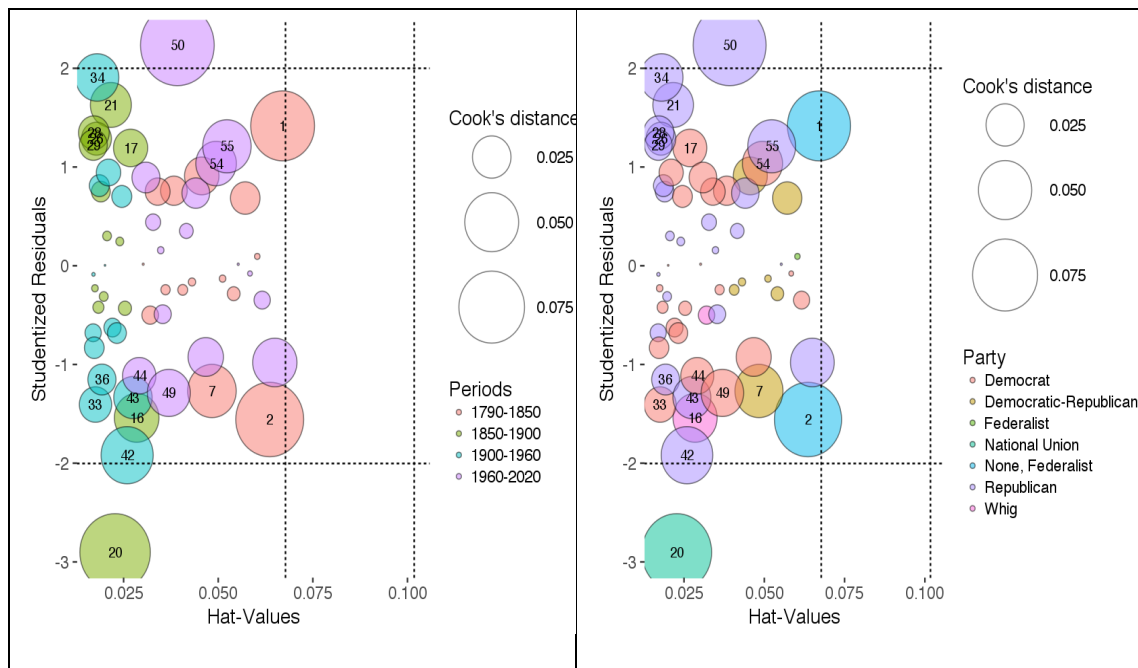


Figure 10: Trump Influence Plot Parties and Periods

Our method has thus identified Eisenhower as the singularity against which Trump sets himself. While this might not be surprising given Eisenhower's global outlook, it is nevertheless surprising that of all the presidents before Trump to describe his discontinuities, we found a Republican and if we include Lincoln even two Republicans. Through attention to details and singularities, our analysis has found that Trump is much more indicative of a Republican struggle with their own ideas rather than struggles with the Democrats.

As the visualization of Cook's Distance can be difficult to read in the influence plots, let us again plot a simple bar plot to identify the three most important influences on the centre of Trump's relationship with this predecessors' speeches. In Figure 11, we can ignore the two first Inaugurals, as we control for leverage, and see that Eisenhower is the strongest negative influence on Trump that is closest to him in time. Also remarkable is the strong

(positive) influence of Reagan and the strong (negative) influence of Lincoln. Trump's positioning as dissimilar to Eisenhower but also overly similar to Reagan sets him within a particular history of conservatism, which is a history of struggle internal to the Republican party, rather than as a "weird" anomaly himself.

While struggles in the Republican party are not surprising, our analyses show – counter-intuitively through anomaly detection and influence tracking – that Trump himself is not a discontinuity from the US presidential 'tradition', but symbolizes a particular struggle within the Republican party. As Colin Koopman has argued, Foucault has moved in his analyses from history as rupture to history as "continuity-with-discontinuity" (Koopman 2013, 42). Our methodological experimentation with multiple temporalities has allowed us to trace the entanglement between continuity and discontinuity rather than privilege continuity and see discontinuity as derivative, or simply privilege discontinuity as the writing of unrelated breaks.

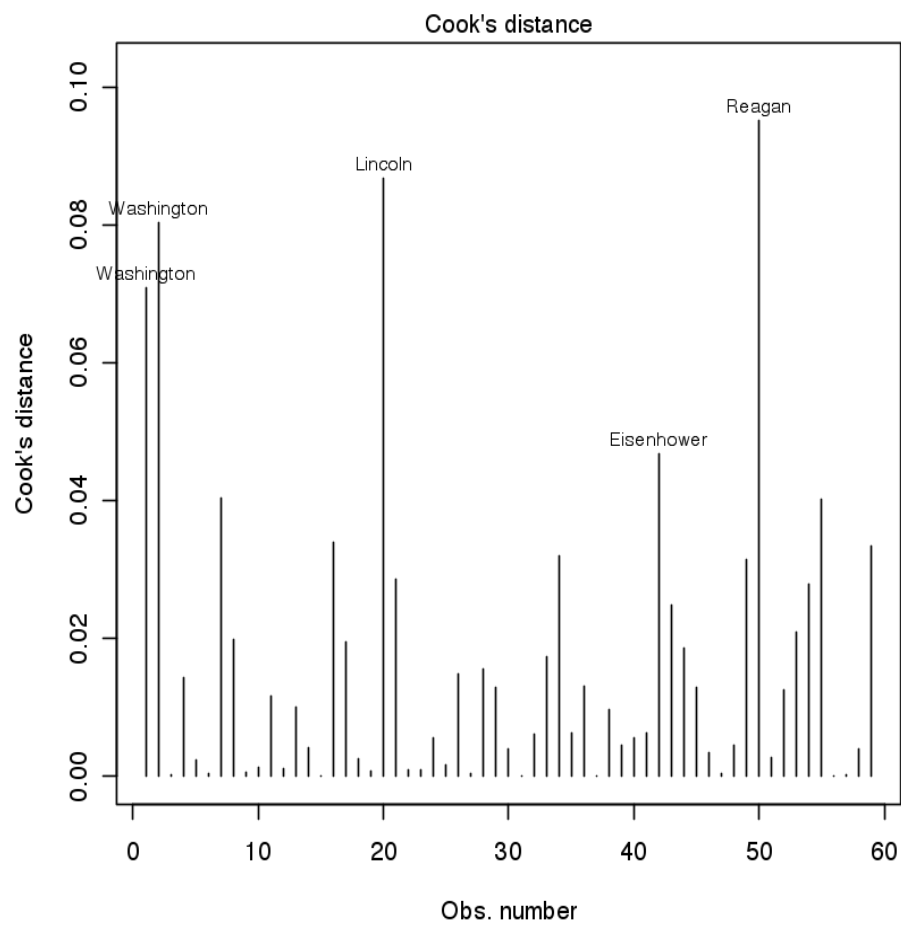


Figure 11: Cook's Distance Trump

## Conclusion

This article has proposed a novel approach to computational historical research, which we have called computational genealogy. We started by placing difference, disruption and discontinuity at the centre of a critical historical analysis and built on Foucault's genealogical methods of attending to dispersion and discontinuities. We follow Koopman's (2013) articulation of genealogy as engagement with complex temporalities where continuity and discontinuity are entangled. Our proposal for computational genealogy does not mean that our methods simply equate with the Foucault's genealogy or that they 'upgrade' it for the digital age. Rather, the computational genealogy draws on a range of different methods and thus introduces new vocabularies of discontinuity as they appear in computational analysis: anomalies, spikes, outliers, influence, de-trending, etc.

As Foucault had noted, discontinuity takes a variety of forms and we need more vocabularies to attend empirically and conceptually to discontinuity and its entanglements with other temporalities (trends, series, chronologies, timelines). Yet, our method builds on the critical sensibility of genealogy to offer an understanding of the contingency of the present and avoid either assumptions of unbroken 'tradition' or linear rises. In a small experiment of computational genealogy, we compared Obama's and Trump's inauguration speeches with all other speeches before them and explored various digital methods to discover entangled continuities and discontinuities.

Obama's first Inaugurals discontinues what Ford wanted in his inauguration, who famously delivered a speech not of new departures but one



that “assumed” the presidency without a strong mandate. For Trump, once we determined his rhetorical tradition, it was other Republicans who stood out. Eisenhower expresses a Republican tradition of internationalism, order and collaboration with others based on an American ideal of free nations. Overall, we found Trump’s rhetoric to be distinct not so much from his direct Democrat predecessors, but from Republicans. While Trump might have been perceived as “weird” by his compatriots, his speech highlights a struggle within the Republican party that culminates in his presence and the disappearance of Eisenhower’s ideas. The trend to Trump is most influenced by Reagan’s conservatism and Republican ultra-nationalism. It is interrupted, however, also by Republicans, by Lincoln’s national consolidation and Eisenhower’s internationalism. The computational genealogy of Trump incorporates all these temporalities.

We started by asking how we could compare Trump’s inaugural speech with those of presidents before him to find out how “weird” his speech was. In order to do so, we had to devise a new method to computationally challenge the unity of historical periods and epochs. Our analysis was conducted on a small archive of Inaugural speeches but there are now more and more and more textual collections that span maybe not centuries like the Inaugurals but at least decades. Our method could be applicable to these, too. Moreover, several types of historical collections could be combined to trace different rhetorical struggles and broaden the analysis of discontinuities beyond presidential elite discourses.

## References

- Ali, Y (2017) *What George W. Bush Really Thought of Donald Trump's Inauguration* 2017 [accessed 5/5/2018]. Available from <http://nymag.com/daily/intelligencer/2017/03/what-george-w-bush-really-thought-of-trumps-inauguration.html>.
- Armitage, D, and J Guldi (2015) The History Manifesto: A Reply to Deborah Cohen and Peter Mandler, The History Manifesto: A Reply. *The American Historical Review* vol. 120 (2):543-554.
- Barabak, MZ (2017) *Raw, angry and aggrieved, President Trump's inaugural speech does little to heal political wounds* 2017. [accessed 5/5/2018], Available from <http://www.latimes.com/politics/la-na-pol-trump-inauguration-speech-analysis-20170120-story.html>.
- Casey, ML (2016) *Working as Civic and Patriotic Duty for Consumption: A Critical Discourse Analysis of American Presidential Inaugural Speeches Since World War II*, UNLV Theses, Dissertations, Professional Papers, and Capstones.
- Chandola, V, A Banerjee, and V Kumar (2009) Anomaly detection: A survey. *ACM Comput. Surv.* vol. 41 (3):1-58.
- Cho, S. G., Cho, J. and Kim, S. B. 2015. Discovering Meaningful Trends in the Inaugural Addresses of United States Presidents Via Text Mining *Journal of Korean Institute of Industrial Engineers*. Korean Institute of Industrial Engineers, 41(5): 453–460. doi: 10.7232/jkiie.2015.41.5.453.
- Cohen, D, and P Mandler (2015) The History Manifesto: A Critique The History Manifesto: A Critique. *The American Historical Review* vol. 120 (2):530-542.
- Cowpertwait, PS, and AV Metcalfe (2009) *Introductory time series with R*, New York: Springer Science & Business Media.

Dreyfus, Hubert L., and Paul Rabinow. Michel Foucault: beyond Structuralism and Hermeneutics. Chicago: Chicago University Press, 1983.

Eisenhower, D (1953) *Dwight Eisenhower's First Inaugural Address* 1953. [accessed 5/5/2018], Available from [https://en.wikisource.org/wiki/Dwight Eisenhower%27s First Inaugural Address](https://en.wikisource.org/wiki/Dwight_Eisenhower%27s_First_Inaugural_Address).

Fahrenthold, DA, R Costa, and J Wagner (2017) *Donald Trump is sworn in as president, vows to end 'American carnage'* 2017. [accessed 5/5/2018], Available from [https://www.washingtonpost.com/politics/trump-to-be-sworn-in-marking-a-transformative-shift-in-the-countrys-leadership/2017/01/20/954b9cac-de7d-11e6-ad42-f3375f271c9c\\_story.html](https://www.washingtonpost.com/politics/trump-to-be-sworn-in-marking-a-transformative-shift-in-the-countrys-leadership/2017/01/20/954b9cac-de7d-11e6-ad42-f3375f271c9c_story.html).

Fairbanks, E (2009) *Should Obama's Inaugural Address Be Bleak?* 2009. [accessed 5/5/2018], Available from <https://newrepublic.com/article/47084/should-obamas-inaugural-address-be-bleak>.

Ford, G (1974) *Gerald Ford's assumption of the Presidency* 1974. [accessed 5/5/2018], Available from [https://en.wikisource.org/wiki/Gerald Ford%27s assumption of the Presidency](https://en.wikisource.org/wiki/Gerald_Ford%27s_assumption_of_the_Presidency).

Foucault, M (1980) *Power/knowledge: Selected interviews and other writings, 1972-1977*, New York: Pantheon.

Foucault, M (1984) Nietzsche, genealogy, history, in: Rabinow, P, *The Foucault Reader*, New York: Pantheon, 76-100.

Foucault, M and Ewald, F. (2003) *Society Must Be Defended, Lectures at the Collège de France, 1975-1976*. London: Macmillan.

Foucault, M (2007) *Security, territory, population: lectures at the Collège de France, 1977-78*, New York: Springer.

Foucault, M (2013) *Archaeology of knowledge*. London: Routledge.

- Fox, J (2009) *Regression Diagnostics* 2009 [cited 1 December 2017]. [accessed 5/5/2018], Available from <https://socialsciences.mcmaster.ca/jfox/Courses/Brazil-2009/index.html>.
- Fuller, WA (2009) *Introduction to statistical time series*. Hoboken: John Wiley & Sons.
- Harding, W (1921) *Warren Harding's Inaugural Address* 1921. [accessed 5/5/2018], Available from [https://en.wikisource.org/wiki/Warren\\_Harding%27s\\_Inaugural\\_Address](https://en.wikisource.org/wiki/Warren_Harding%27s_Inaugural_Address).
- Harding, WG (1920) *Return to Normalcy* 1920. [accessed 5/5/2018], Available from <http://teachingamericanhistory.org/library/document/return-to-normalcy/>.
- Hayles, NK (2017) *Unthought: The Power of the Cognitive Nonconscious*. Chicago: University of Chicago Press.
- Kejariwal, A (2015) *Introducing practical and robust anomaly detection in a time series* 2015. [accessed 5/5/2018], Available from <https://blog.twitter.com/2015/introducing-practical-and-robust-anomaly-detection-in-a-time-series>.
- Koopman, C (2013) *Genealogy as critique: Foucault and the problems of modernity*. Bloomington: Indiana University Press.
- LaCapra, Dominick. *History and Reading: Tocqueville, Foucault, French Studies*. Toronto: University of Toronto Press, 2000.
- Lansdall-Welfare, T, et al. (2017) Content analysis of 150 years of British periodicals. *Proceedings of the National Academy of Sciences* vol. 114 (4): E457-E465.
- Light, R (2014) From Words to Networks and Back. *Social Currents* vol. 1 (2):111-129.
- Lim, ET (2002) Five trends in presidential rhetoric: An analysis of rhetoric from George Washington to Bill Clinton. *Presidential Studies Quarterly* vol. 32 (2): 328-348.
- Lincoln, A (1965) *Abraham Lincoln's Second Inaugural Address* 1965. [accessed 5/5/2018], Available from [https://en.wikisource.org/wiki/Abraham\\_Lincoln%27s\\_Second\\_Inaugural Address](https://en.wikisource.org/wiki/Abraham_Lincoln%27s_Second_Inaugural_Address).

- Michel, J-B, et al. (2011) Quantitative analysis of culture using millions of digitized books. *science* vol. 331 (6014): 176-182.
- Obama, B. (2009) Barak Obama's first inaugural address. [accessed 5/5/2018], Available from [https://en.wikisource.org/wiki/Barack\\_Obama%27s\\_First\\_Inaugural\\_Address](https://en.wikisource.org/wiki/Barack_Obama%27s_First_Inaugural_Address)
- Pierce, F (1853) Franklin Pierce's Inaugural Address 2017. [accessed 5/5/2018], Available from [https://en.wikisource.org/wiki/Franklin\\_Pierce%27s\\_Inaugural\\_Address](https://en.wikisource.org/wiki/Franklin_Pierce%27s_Inaugural_Address).
- Prabhakaran, S (2016) *Outlier Treatment* 2016. [accessed 5/5/2018], Available from <http://r-statistics.co/Outlier-Treatment-With-R.html>.
- Reagan, R. 1981. Ronald Reagan's inaugural address. [accessed 5/5/2018], [https://en.wikisource.org/wiki/Ronald\\_Reagan%27s\\_First\\_Inaugural\\_Address](https://en.wikisource.org/wiki/Ronald_Reagan%27s_First_Inaugural_Address).
- Robin, C (2011) *The Reactionary Mind: Conservatism from Edmund Burke to Sarah Palin*. Oxford: Oxford University Press.
- Rogers, R (2013) *Digital methods*. Cambridge, Massachusetts: MIT press.
- Rosner, B (1983) Percentage points for a generalized ESD many-outlier procedure. *Technometrics* vol. 25 (2): 165-172.
- Rule, A, J-P Cointet, and PS Bearman (2015) Lexical shifts, substantive changes, and continuity in State of the Union discourse, 1790–2014. *Proceedings of the National Academy of Sciences* vol. 112 (35): 10837-10844.
- Trump, D. (2016), Donald Trump's inaugural address. [accessed 5/5/2018], [https://en.wikisource.org/wiki/Donald\\_Trump%27s\\_Inaugural\\_Address](https://en.wikisource.org/wiki/Donald_Trump%27s_Inaugural_Address).
- Vallis, O, J Hochenbaum, and A Kejariwal (2014) A Novel Technique for Long-Term Anomaly Detection in the Cloud. Paper read at 6th USENIX Workshop on Hot Topics in Cloud Computing (HotCloud 14), Philadelphia, PA
- Veyne, P. (1997), Foucault revolutionizes history. In Arnold Ira Davidson (ed.), *Foucault and His Interlocutors*. Chicago: University of Chicago Press. pp. 146--82.

Weisberg, S (2005) *Applied linear regression*. Hoboken: John Wiley & Sons.

Whissell, C, and L Sigelman (2001) The Times and the Man as Predictors of Emotion and Style in the Inaugural Addresses of U.S. Presidents. *Computers and the Humanities* vol. 35 (3): 255-272.

Wikipedia (2017) *U.S. Presidential Inaugural Addresses* 2017. [accessed 5/5/2018], Available from [https://en.wikisource.org/wiki/Category:U.S. Presidential Inaugural Addresses](https://en.wikisource.org/wiki/Category:U.S._Presidential_Inaugural_Addresses).

---

<sup>i</sup> See also The American Presidency Project: <http://www.presidency.ucsb.edu/>

<sup>ii</sup> Please note we use  $4/n$  as a cut-off value for Cook's distance, where  $n$  is the number of speeches.